



SEQUENCE LISTING

<110> Genencor International, Inc.
Jones, Brian E.
Grant, William D.
Heaphy, Shaun
Grant, Susan
Rees, Helen

<120> Novel Bacillus mHKcel Cellulase

<130> GC797-2-PCT

<140> PCT/US2004/013257

<141> 2004-04-28

<150> US 60/467,315

<151> 2003-04-30

<160> 3

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 3796

<212> DNA

<213> Bacillus sp.

<220>

<221> misc_feature

<222> (1)...(3796)

<223> isolated from environmental water sample from
Sonachi Lake, Kenya

<221> misc_feature

<222> (1)...(3796)

<223> n = A,T,C or G

<400> 1

atctttagag	atgagcgtga	tagtcagtgg	accttaaaaag	aacttaaaaag	atttttagaa	60
gagattaaaa	cggacccgca	ccatctctct	gtttattttg	atgggggatt	tgatttggag	120
acacaacgat	ctggccttgg	gtgtgtgatt	tattatgaac	aaaatgacac	gtcttaccgg	180
gtgagaaaaa	acgctaccgt	ggcgtcattg	acatcgaata	accaagcaga	atatgccgct	240
ttgcatttag	gacttaaaaag	aacttgaagg	gatcgggcgc	atcatctgcc	tatcaccatt	300
tacggtgatt	ctcaagttgt	gatcaatcag	ttaaaaggag	aatgggcccgt	gtatggagga	360
ggtgttaaat	aaatgggctg	acccgtattg	atcaagcatt	tagctaaatt	aggcatgacc	420
gctacttata	agttaatccc	ccgtaaagaa	aaccgcgaag	ccgatcaact	ggctacacaa	480
gcgttaaacg	ggcaagaaat	tataagtcaa	cgtgatatca	gtgagcgtgg	tgcagattag	540
gctgcaccgc	gcaaaaaaag	tcaacgtgtt	taggaatgga	cagggattaa	agcaacataa	600
ttctctctaa	gcaaacgttg	cgacagcaag	agagaagcat	ataaggtttt	tctgagttag	660
tctatttata	ccaatgtcca	cgtactaaat	aaacctctca	tcaaagtgga	ttttttgatt	720
aattcacttc	cactcctacc	tttatctata	taaattagtt	ccttttttgt	taataatcac	780
taattttggc	ggtatttttt	aatagaaata	tatgctagat	tataaactag	tagccgtata	840
gaagggtggtg	attgccccta	taagagacgt	ctggcaaaca	taaaagcatc	gcattattat	900
aatcgaaaagg	tggagatgag	acatgggtta	tacccaagct	aagtgtatgg	tgaaaaaaac	960

ggtcttgttt	ggtttaattc	tctgtttagg	tgtgtcaatg	tttgtaccag	ttacatcagc	1020
tgaagatagg	gtctcttcgt	cacagggtgga	tatccaatca	tatgtagcag	atatgcaacc	1080
tggctggaat	ttaggtaata	catttgatgc	gataggagat	gatgaaacag	catggggaaa	1140
ccctcgtgta	acgagagaat	taatagaaat	gattgctgat	gaagggtata	aaagtattcg	1200
tatcccagtc	acatggcaaa	atcaaattggg	tggttctcca	gattatacaa	ttaatgaaga	1260
ttatatcaag	cgggtagagc	aagtgataga	ttgggcgttg	gaggaagact	tgtatgtgat	1320
gttaaattgtg	catcatgact	catggctgtg	gatgtatgat	atggaacata	actatgatga	1380
ggtgatggca	agatatacag	ctatttggga	acaattgtcg	gaaaaattca	aaaaccactc	1440
ccataagttg	atgtttgaga	gtgtcaatga	gcctaggttt	acgcaggagt	ggggagagat	1500
tcaagaaaat	catcatgctt	acttagaaga	tttaaataag	acgttctatt	atattgtcag	1560
agagtcagga	ggcaataatg	tggagcgccc	tttagtattg	cctacgatag	aaacagccac	1620
gtctcaggat	ttactagatc	gcttgtatca	aacaatggaa	gacttggatg	accctcattt	1680
aattgccacg	gttcattatt	atggcttttg	gccccttagt	gtcaatatag	cagggtacac	1740
ccgttttgaa	caggagacac	aacaagatat	tatagacacg	tttgaccgtg	ttcataacac	1800
atttacagcg	aatgggatcc	cagttgtatt	aggtgaattt	ggtttgttag	gctttgataa	1860
aagtacggac	gtcattcagc	aaggtagaaa	attaaaattt	tttgagtttc	tcatccatca	1920
tctcaatgaa	cgtgatataa	cccatatgtt	atgggataac	ggtcagcatt	taaagcgaga	1980
aacttattca	tggtagatc	aggaatttca	tgacatatta	aaagcgagtt	gggaggggcg	2040
ttctgctaca	gctgagtcta	atttcattca	tgtgaaggac	ggagagccaa	ttagagatca	2100
acataatacag	ctttacttaa	acggaaatga	gctaactgcc	ctacaggcag	gggacgaatc	2160
gcttgacta	ggagaggatt	atgagctagc	aggagacgta	ttaacgctaa	aagcgggcat	2220
cctcacaaga	ttaattaccc	ctggccaatt	aggaacgaat	gcggtcatca	cagctcaatt	2280
taattctgga	gcagactggc	gttttcaatt	acagaatgtg	gacgtgccaa	cagtcgaaaa	2340
tacagatggc	tcaatatggc	attttgcgat	ccctacccat	tttaatggtg	atagtcttgc	2400
gacgatggaa	gctgtttatg	caaacggaga	atatgctggc	ccgcaagatt	ggacgtcatt	2460
taaagaattt	ggcgaggcgt	tttcccctaa	ttacgccaca	ggggaaatta	ttataacaga	2520
agccttcttt	aacgcggtag	gggatgatga	tatccattta	acatttcatt	attggagcgg	2580
agagacgggtg	gaatatacat	tacgtaaaaa	tggaaattat	gttcaaggta	gacggtaaaa	2640
tgattttta	taatatagata	aacagcctac	ctatcgtttt	tggagaagg	caaacgaatc	2700
tcatcttacc	aacaccgtgc	tttagaactt	tagaagtgc	ggtgtttttt	aagacatgag	2760
gagagacaat	cctctatcaa	cagtcaccaa	tttttattca	ggaggtgtca	agttatctaa	2820
cgttctatga	atgcataatg	tttctgacga	ataaacatag	ttaaaaagaa	gtgagcctag	2880
ttcccaggagg	gaaggggata	atgccaacgt	attggattaa	agtaccttct	tgataaaaaag	2940
aaagggtttt	caagagggtg	aatgggctc	gtttgttata	ctttaattac	accttgggac	3000
gtcatttttg	cgggtgctact	tagtaagatg	actgacatca	taaaagagga	gtgggttcga	3060
tggctttaat	tcaattaagc	tttaaatcac	gagcattaat	gttgcaaacc	tctgtcaatg	3120
ttttattacc	ggtgggaatg	aatgcggtag	attttacacc	aagtgatgat	ttttcttatg	3180
ttactgaccc	ttttcctgtc	ctatatcttt	tgcatggtgc	aactgatgat	tattcagcat	3240
ggctacgtct	gtcctccatt	gaacgatatg	ctgaagaaaa	aaaattggcg	gtcgtcatgc	3300
caaatgctga	tatgagtgcg	tatacggata	tggtagatgg	acatcgttac	tggacgtata	3360
ttagtaagga	gctgcctgag	tttatcaaag	cgacttttcc	tatttctcag	caccgtgaag	3420
acacctttgc	ggctggctctg	tctatgggag	gatacggggc	ttttaaatga	gcgttgccgc	3480
aaccggaacg	cttcgctgca	gntgtgtcat	tatcagggtgc	agttgatatg	agagaagcaa	3540
gtcaaccaga	ctccctattt	gtgaacgcat	ttggtgaagg	gacgaaaatc	gcagggacag	3600
atcttgatct	ttttcattta	attaaaaagt	tgggggtata	tgaaggggct	aaaccagccc	3660
tttttcaagc	gtgtgggaca	gaggactttt	tatatgaaga	taatgtgaga	tttagagatt	3720
atgcacgaca	agtgaatgcc	gatttaactt	atgaagaagg	tcctggtggt	catgaatggg	3780
cttattggga	tagaat					3796

<210> 2

<211> 1716

<212> DNA

<213> Bacillus sp.

<220>

<221> misc_feature

<222> (1)...(1716)

<223> isolated from environmental water sample from
Sonachi Lake, Kenya

<400> 2

atggggttata	cccaagctaa	gtgtatggtg	aaaaaaacgg	tcttgtttg	tttaattctc	60
tgtttaggtg	tgtcaatggt	tgtaccagtt	acatcagctg	aagatagggt	ctcttcgtca	120
caggtggata	tccaatcata	tgtagcagat	atgcaacctg	gctggaattt	aggtaataca	180
tttgatgcga	taggagatga	tgaacacagca	tggggaaacc	ctcgtgtaac	gagagaatta	240
atagaaatga	ttgctgatga	aggggtataaa	agtattcgta	tcccagtcac	atggcaaaat	300
caaatgggtg	gttctccaga	ttatacaatt	aatgaagatt	atatcaagcg	ggtagagcaa	360
gtgatagatt	gggcgttgga	ggaagacttg	tatgtgatgt	taaatgtgca	tcatgactca	420
tggctgtgga	tgtatgatat	ggaacataac	tatgatgagg	tgatggcaag	atatacagct	480
atgttggaac	aattgtcggg	aaaattcaaa	aaccactccc	ataagttgat	gtttgagagt	540
gtcaatgagc	ctaggtttac	gcaggagtgg	ggagagattc	aagaaaatca	tcatgcttac	600
ttagaagatt	taaataagac	gttctattat	attgtcagag	agtcaggagg	caataatgtg	660
gagcgccctt	tagtattgcc	tacgatagaa	acagccacgt	ctcaggattt	actagatcgc	720
ttgtatcaaa	caatggaaga	cttggtatgac	cctcatttaa	ttgccacggg	tcattattat	780
ggcttttggc	ccttttagtgt	caatatagca	gggtacaccc	gttttgaaca	ggagacacaa	840
caagatatta	tagacacggt	tgaccgtggt	cataacacat	ttacagcgaa	tgggatccca	900
gttgatttag	gtgaatttgg	tttgtttaggc	tttgataaaa	gtacggacgt	cattcagcaa	960
ggtgagaaat	taaaattttt	tgagtttctc	atccatcatc	tcaatgaacg	tgatataacc	1020
catatgttat	gggataacgg	tcagcattta	aagcgagaaa	cttattcatg	gtatgatcag	1080
gaatttcatt	acataattaa	agcgagttgg	gaggggctgt	ctgctacagc	tgagtctaata	1140
ttcattcatg	tgaaggacgg	agagccaatt	agagatcaac	atatacagct	ttacttaaac	1200
ggaaatgagc	taactgccct	acaggcaggg	gacgaatcgc	ttgtactagg	agaggattat	1260
gagctagcag	gagacgtatt	aacgctaata	gcgggcatcc	tcacaagatt	aattaccctt	1320
ggccaattag	gaacgaatgc	ggtcatcaca	gctcaattta	attctggagc	agactggcgt	1380
tttcaattac	agaatgtgga	cgtgccaaca	gtcgaaaata	cagatggctc	aatatggcat	1440
tttgcatccc	ctaccatttt	taatgggtgat	agtcttgcca	cgatgggaagc	tgtttatgca	1500
aacggagaat	atgctggccc	gcaagattgg	acgtcattta	aagaatttgg	cgaggcgttt	1560
ttccctaatt	acgccacagg	ggaaattatt	ataacagaag	ccttctttta	cgcggtacgg	1620
gatgatgata	tccatttaac	atttcattat	tggagcggag	agacggtgga	atatacatta	1680
cgtaaaaatg	gaaattatgt	tcaaggtaga	cggttaa			1716

<210> 3

<211> 571

<212> PRT

<213> Bacillus sp.

<220>

<221> VARIANT

<222> (1)...(571)

<223> isolated from environmental water sample from
Sonachi Lake, Kenya

<400> 3

Met	Gly	Tyr	Thr	Gln	Ala	Lys	Cys	Met	Val	Lys	Lys	Thr	Val	Leu	Phe
1				5					10					15	
Gly	Leu	Ile	Leu	Cys	Leu	Gly	Val	Ser	Met	Phe	Val	Pro	Val	Thr	Ser
			20					25					30		
Ala	Glu	Asp	Arg	Val	Ser	Ser	Ser	Gln	Val	Asp	Ile	Gln	Ser	Tyr	Val
			35				40					45			
Ala	Asp	Met	Gln	Pro	Gly	Trp	Asn	Leu	Gly	Asn	Thr	Phe	Asp	Ala	Ile
			50				55				60				
Gly	Asp	Asp	Glu	Thr	Ala	Trp	Gly	Asn	Pro	Arg	Val	Thr	Arg	Glu	Leu
65					70				75					80	
Ile	Glu	Met	Ile	Ala	Asp	Glu	Gly	Tyr	Lys	Ser	Ile	Arg	Ile	Pro	Val

										85					90					95				
Thr	Trp	Gln	Asn	Gln	Met	Gly	Gly	Ser	Pro	Asp	Tyr	Thr	Ile	Asn	Glu									
			100					105					110											
Asp	Tyr	Ile	Lys	Arg	Val	Glu	Gln	Val	Ile	Asp	Trp	Ala	Leu	Glu	Glu									
			115					120					125											
Asp	Leu	Tyr	Val	Met	Leu	Asn	Val	His	His	Asp	Ser	Trp	Leu	Trp	Met									
			130					135					140											
Tyr	Asp	Met	Glu	His	Asn	Tyr	Asp	Glu	Val	Met	Ala	Arg	Tyr	Thr	Ala									
145					150					155					160									
Ile	Trp	Glu	Gln	Leu	Ser	Glu	Lys	Phe	Lys	Asn	His	Ser	His	Lys	Leu									
			165					170					175											
Met	Phe	Glu	Ser	Val	Asn	Glu	Pro	Arg	Phe	Thr	Gln	Glu	Trp	Gly	Glu									
			180					185					190											
Ile	Gln	Glu	Asn	His	His	Ala	Tyr	Leu	Glu	Asp	Leu	Asn	Lys	Thr	Phe									
			195					200					205											
Tyr	Tyr	Ile	Val	Arg	Glu	Ser	Gly	Gly	Asn	Asn	Val	Glu	Arg	Pro	Leu									
210					215					220					225									
Val	Leu	Pro	Thr	Ile	Glu	Thr	Ala	Thr	Ser	Gln	Asp	Leu	Leu	Asp	Arg									
			230					235					240											
Leu	Tyr	Gln	Thr	Met	Glu	Asp	Leu	Asp	Asp	Pro	His	Leu	Ile	Ala	Thr									
			245					250					255											
Val	His	Tyr	Tyr	Gly	Phe	Trp	Pro	Phe	Ser	Val	Asn	Ile	Ala	Gly	Tyr									
			260					265					270											
Thr	Arg	Phe	Glu	Gln	Glu	Thr	Gln	Gln	Asp	Ile	Ile	Asp	Thr	Phe	Asp									
			275					280					285											
Arg	Val	His	Asn	Thr	Phe	Thr	Ala	Asn	Gly	Ile	Pro	Val	Val	Leu	Gly									
290					295					300					305									
Glu	Phe	Gly	Leu	Leu	Gly	Phe	Asp	Lys	Ser	Thr	Asp	Val	Ile	Gln	Gln									
			310					315					320											
Gly	Glu	Lys	Leu	Lys	Phe	Phe	Glu	Phe	Leu	Ile	His	His	Leu	Asn	Glu									
			325					330					335											
Arg	Asp	Ile	Thr	His	Met	Leu	Trp	Asp	Asn	Gly	Gln	His	Leu	Lys	Arg									
			340					345					350											
Glu	Thr	Tyr	Ser	Trp	Tyr	Asp	Gln	Glu	Phe	His	Asp	Ile	Leu	Lys	Ala									
			355					360					365											
Ser	Trp	Glu	Gly	Arg	Ser	Ala	Thr	Ala	Glu	Ser	Asn	Phe	Ile	His	Val									
370					375					380					385									
Lys	Asp	Gly	Glu	Pro	Ile	Arg	Asp	Gln	His	Ile	Gln	Leu	Tyr	Leu	Asn									
			390					395					400											
Gly	Asn	Glu	Leu	Thr	Ala	Leu	Gln	Ala	Gly	Asp	Glu	Ser	Leu	Val	Leu									
			405					410					415											
Gly	Glu	Asp	Tyr	Glu	Leu	Ala	Gly	Asp	Val	Leu	Thr	Leu	Lys	Ala	Gly									
			420					425					430											

His Leu Thr Phe His Tyr Trp Ser Gly Glu Thr Val Glu Tyr Thr Leu
545 550 555 560
Arg Lys Asn Gly Asn Tyr Val Gln Gly Arg Arg
565 570